Available online at www.sciencedirect.com
ScienceDirect
Discrete Mathematics 308 (2008) 6009–6014
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

Optimality conditions for Hunter's bound

Pierangela Veneziani

Department of Mathematics, SUNY College at Brockport, 350 New Campus Drive, Brockport, NY 14420, USA

Received 24 January 2006; received in revised form 6 March 2007; accepted 12 November 2007
Available online 26 December 2007

Abstract

The bound known as Hunter's bound states that $P(A_1 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} p_i - \sum_{\{i,j\} \in T} p_{i,j}$, where T designates the heaviest spanning tree of the graph on n nodes with edge weights $p_{i,j}$. We prove that Hunter's bound is optimal if and only if the input probabilities are given on a tree.

Keywords: Hunter's bound; Probability bounds; Probability of the union

1. Introduction

The *Boolean probability bounding problem* can be formulated as follows.

Let $A_1, \ldots, A_n$ be a finite set of arbitrary events in a probability space Ω, and let us assume that the individual probabilities $P(A_i)$, $i = 1, \ldots, n$, as well as the probabilities $P\left(\bigcap_{1 \leq i_1 < \cdots < i_l \leq n} A_{i_l}\right)$, $l = 2, \ldots, m$, up to m-tuples of these events are known, where $m < n$. Using this information we want to generate upper and lower bounds for the probability of a Boolean function of these events. The integer m is usually referred to as the *degree* of these bounds.

Let us introduce the following notations: let $\mathcal{G}_m = (\mathcal{V}, \mathcal{E})$ denote the hypergraph, where $\mathcal{V} = \{1, \ldots, n\}$ and $\mathcal{E} = \bigcup_{k=2}^{m} \mathcal{E}_k$, where $\mathcal{E}_k = \{I \subseteq \mathcal{V} \mid |I| = k\}$, $k = 2, \ldots, m$. Further let $\Gamma = \mathcal{V} \cup \mathcal{E}$.

For each subset $J \subseteq \mathcal{V}$ let us define the event $C_J = \left(\bigcap_{i \in J} A_i\right)\left(\bigcap_{i \in J^c} A_i^c\right)$, where $J^c = \mathcal{V} \setminus J$, and $A_i^c = \Omega \setminus A_i$, $i = 1, \ldots, n$, and to each subset $J \subseteq \mathcal{V}$ let us associate a decision variable $x_J = \Pr(C_J)$ and a scalar c_J.

Let us further introduce the notation $p_I = P\left(\bigcap_{i \in I} A_i\right)$, where $I \in \Gamma$, and let us set $p_\emptyset = 1$ by definition.

Let us note that the equality $\sum_{I \subseteq J} P(C_J) = P\left(\bigcap_{i \in I} A_i\right)$ holds for all subsets $I \in \Gamma \cup \{\emptyset\}$, because the 2^n (disjoint) events C_J's form a partition of the probability space Ω. We can then write the last equality as $\sum_{I \subseteq J \subseteq \mathcal{V}} x_J = p_I$.

Finally, let p denote the vector with components $p_I \in [0, 1]$, $I \in \Gamma \cup \{\emptyset\}$, let x be the vector with components $x_J \in [0, 1]$, $J \subseteq \mathcal{V}$, and let $H = (h_{IJ})$ denote the incidence matrix whose entries are defined by

$$h_{IJ} = \begin{cases} 1 & \text{if } I \subseteq J \\ 0 & \text{otherwise.} \end{cases}$$

The matrix H has $\sum_{i=0}^{m} \binom{n}{i}$ rows and 2^n columns.

E-mail address: pvenezia@brockport.edu.

doi:10.1016/j.disc.2007.11.019

In the vectors p and x, and in the row and column indices of the matrix H the order of the elements will follow the lexicographic order of the subscript sets.

The Boolean probability bounding problem can thus be restated as a linear program of the form

$$\begin{array}{lll} \text{Max or Min} & \sum_{J\subseteq\mathcal{V}} c_J x_J & \\ \text{st} & \sum_{I\subseteq J\subseteq\mathcal{V}} h_{IJ} x_J = p_I & \forall I \in \Gamma \cup \{\emptyset\} \\ & x_J \geq 0 & \forall J \subseteq \mathcal{V} \end{array}$$

or in matrix form as

$$\begin{array}{ll} \text{Max} & c^{\mathrm{T}} x \\ \text{st} & Hx = p \\ & x \geq 0 \end{array} \tag{1}$$

and

$$\begin{array}{ll} \text{Min} & c^{\mathrm{T}} x \\ \text{st} & Hx = p \\ & x \geq 0, \end{array} \tag{2}$$

where the vector c has components c_J, $J \subseteq \mathcal{V}$.

In particular, if $c^{\mathrm{T}} = [0, 1, \ldots, 1]$, problems (1) and (2) provide us with bounds for the probability $P(A_1 \cup \cdots \cup A_n)$ that at least one out of n events occurs.

As an illustration consider for example the case $n = 3$, $m = 2$, $c^{\mathrm{T}} = [0, 1, \ldots, 1]$, $p_I = 0.5$ for $|I| = 1$, $p_I = 0.25$ for $|I| = 2$:

$$\begin{array}{lcccccccl} \text{Max} & & x_1 + x_2 + x_3 + x_{12} + x_{13} + x_{23} + x_{123} \\ \text{st} \\ x_\emptyset & +x_1 & +x_2 & +x_3 & +x_{12} & +x_{13} & +x_{23} & +x_{123} & = 1 \\ & x_1 & & & +x_{12} & +x_{13} & & +x_{123} & = 0.5 \\ & & x_2 & & +x_{12} & & +x_{23} & +x_{123} & = 0.5 \\ & & & x_3 & & +x_{13} & +x_{23} & +x_{123} & = 0.5 \\ & & & & x_{12} & & & +x_{123} & = 0.25 \\ & & & & & x_{13} & & +x_{123} & = 0.25 \\ & & & & & & x_{23} & +x_{123} & = 0.25 \\ & & x_\emptyset, x_1, x_2, x_3, x_{12}, x_{13}, x_{23}, x_{123} \geq 0. \end{array}$$

In this example the optimal objective function value of the maximization problem is 1, achieved for $x_\emptyset = x_{12} = x_{13} = x_{23} = 0$ and $x_1 = x_2 = x_3 = x_{123} = 0.25$.

The upper and lower bounds on the linear functional $c^{\mathrm{T}} x$ can be obtained by computing the optimum value of linear programming problems (1) and (2) respectively. Unfortunately the number of variables in (1) and (2) increases exponentially with the number of events, which makes their solution impractical in reasonable time.

Consider then the duals of problems (1) and (2):

$$\begin{array}{ll} \text{Min} & p^{\mathrm{T}} w \\ \text{st} & H^{\mathrm{T}} w \geq c \end{array} \tag{3}$$

and

$$\begin{array}{ll} \text{Max} & p^{\mathrm{T}} w \\ \text{st} & H^{\mathrm{T}} w \leq c. \end{array} \tag{4}$$

Recall that if a linear programming problem is a maximization (minimization), the objective function value corresponding to any dual feasible basis is an upper (lower) bound for its optimum value. The best bound corresponds to the optimal basis and is called sharp because no better bound can be given based on the knowledge of the vector p. Thus, bounds can be obtained provided that we can construct dual feasible bases.

2. Upper bounds of degree 2

Consider problems (1) and (2) for $m = 2$ and cost coefficients $c^{\mathrm{T}} = [0, 1, \ldots, 1]$. The objective function then becomes $\sum_{J\subseteq\mathcal{V}} c_J x_J = \sum_{\emptyset\neq J\subseteq\mathcal{V}} x_J$.

In the linear programming problems (1) and (2) we have $1 + n + \binom{n}{2}$ constraints and 2^n variables.

The first constraint $\sum_{J\subseteq\mathcal{V}} x_J = 1$ becomes superfluous because we are going to maximize and minimize the quantity $\sum_{\emptyset\neq J\subseteq\mathcal{V}} x_J$. The optimum value of the minimization problem is less than or equal to 1 by construction, while if the optimum value of the maximization problem is found to be larger than 1 then, by taking into account the constraint $\sum_{J\subseteq\mathcal{V}} x_J = 1$, we can trivially set the upper bound to 1. Therefore the first row of the matrix H as well as the first column corresponding to the variable $x_\emptyset$ can be disregarded from our formulation. In the linear programming problems (1) and (2) we now have $n + \binom{n}{2}$ constraints and $2^n - 1$ variables.

As Prékopa et al. suggested in [3], it is then possible to interpret the $n + \binom{n}{2}$ components of any dual feasible solution $w = (w_\gamma)_{\gamma\in\Gamma}$ of problems (3) and (4) as nodes and edge weights in $\mathcal{G}_2$, that is a weight w_i is assigned to node $i \in \mathcal{V}$ and a weight $w_{i,j}$ is assigned to edge $\{i, j\} \in \mathcal{E}_2$.

In what follows we will let $\mathcal{E}(S)$ denote the edge set of a subset $S \subseteq \mathcal{V}$ and

$$w(S) = \sum_{\gamma\in S} w_\gamma + \sum_{\gamma\in\mathcal{E}(S)} w_\gamma$$

represent the weight of subset S for a given dual feasible solution $w = (w_\gamma)_{\gamma\in\Gamma}$.

For the instance under study ($c^{\mathrm{T}} = [0, 1, \ldots, 1]$ and $m = 2$) problems (3) and (4) can then be written as

$$\begin{aligned} &\text{Min} \quad \sum_{\gamma\in\Gamma} p_\gamma w_\gamma \\ &\text{st} \qquad w(S) \geq 1 \quad \forall S \subseteq \mathcal{V} \end{aligned} \tag{5}$$

and

$$\begin{aligned} &\text{Max} \quad \sum_{\gamma\in\Gamma} p_\gamma w_\gamma \\ &\text{st} \qquad w(S) \leq 1 \quad \forall S \subseteq \mathcal{V}. \end{aligned} \tag{6}$$

The lemma that follows provides a sufficient and necessary condition for a given vector to be a basic feasible solution of problem (5) by use of the graph structure introduced at the beginning of Section 1.

Lemma 1. *Given a collection $\Im = \{I_\gamma\}_{\gamma\in\Gamma}$ of column subscripts of the matrix H, a vector $w = (w_\gamma)_{\gamma\in\Gamma}$ is a basic feasible solution of problem (5) generated by the basis $\Im$ if the following conditions are satisfied.*

(i) *The vector $w = (w_\gamma)_{\gamma\in\Gamma}$ is the unique solution of the system of equations $w(I_\gamma) = 1$ for all subsets $I_\gamma \in \Im$, $\gamma \in \Gamma$.*

(ii) *For all subsets $S \subseteq \mathcal{V}$ such that $S \notin \Im$ the inequality $w(S) \geq 1$ holds.*

Proof. Let h_J, $J \subseteq \mathcal{V}$, designate a column vector of the matrix H. Let B denote a nonsingular square submatrix of H of order $n + \binom{n}{2}$ and let $\Im = \{I_\gamma\}_{\gamma\in\Gamma}$ denote the collection of subscripts whose columns form B. Recall that a matrix B is said to be a dual feasible basis of problem (1) if $c_B^{\mathrm{T}} B^{-1} h_{I_\gamma} = c_{I_\gamma}$ for all subsets $I_\gamma \in \Im$, $\gamma \in \Gamma$, and $c_B^{\mathrm{T}} B^{-1} h_J \geq c_J$ for all subsets $J \notin \Im$. The corresponding dual basic feasible solution is the vector $w^{\mathrm{T}} = c_B^{\mathrm{T}} B^{-1}$.

In our case, condition (i) guarantees that the matrix B is nonsingular and that the equalities $c_B^{\mathrm{T}} B^{-1} h_{I_\gamma} = c_{I_\gamma}$ hold for all basic sets $I_\gamma \in \Im$, $\gamma \in \Gamma$, and condition (ii) ensures that the inequalities $c_B^{\mathrm{T}} B^{-1} h_J \geq c_J$ are satisfied for all nonbasic sets $J \notin \Im$. $\square$

Remark 2. Let $\mathcal{G}^* = (\Gamma \cup \Im, \mathcal{E}^*)$ denote the bipartite graph, where $\mathcal{E}^* = \{I \in \Gamma, J \in \Im \mid I \subseteq J\}$.

A necessary condition for a collection $\Im = \{I_\gamma\}_{\gamma\in\Gamma}$ of column subscripts of H to form a basis is that there exists a perfect matching in the bipartite graph $\mathcal{G}^*$, otherwise if no perfect matching exists the matrix B would be singular (see e.g. [2]). Therefore in constructing a basis $\Im = \{I_\gamma\}_{\gamma\in\Gamma}$ we want to make sure that we cover all the nodes and edges of $\mathcal{G}_2$.

3. Optimality conditions for Hunter's bound

This section is devoted to the presentation of optimality conditions for the bound known as Hunter's bound [1], which states that

$$P(A_1 \cup \cdots \cup A_n) \le \sum_{i\in\mathcal{V}} p_i - \sum_{\{i,j\}\in T} p_{i,j}, \tag{7}$$

where T designates a spanning tree of the graph $\mathcal{G}_2$. The best bound corresponds to the heaviest spanning tree T.

We will recall [3] that Hunter's bound is generated by minimizing the objective function of problem (5) over the family of dual feasible bases presented in the following proposition.

Proposition 3. *Let T denote a spanning tree of the graph $\mathcal{G}_2$. Then the vector $w = (w_\gamma)_{\gamma\in\Gamma}$ with components*

$$w_\gamma = \begin{cases} 1 & \text{if } \gamma \in \mathcal{V} \\ -1 & \text{if } \gamma \in T \\ 0 & \text{otherwise} \end{cases}$$

is a basic feasible solution of problem (5).

To obtain optimality conditions for Hunter's bound, let us introduce the following Linear Program:

$$\begin{array}{ll} \text{Min} & \sum_{\gamma\in\Gamma_T} p_\gamma w_\gamma = z(w) \\ \text{st} & w_T(S) \ge 1 \quad \forall S \subseteq \mathcal{V}, \end{array} \tag{8}$$

where $\Gamma_T = \mathcal{V} \cup T$ and

$$w_T(S) = \sum_{\gamma\in S} w_\gamma + \sum_{\gamma\in\mathcal{E}(S)\cap T} w_\gamma.$$

Clearly problem (8) is obtained from problem (5) by restricting the node-edge set Γ of the graph $\mathcal{G}_2$ to the node-edge set Γ_T of its subgraph T.

As an illustration consider for example the case $n = 3, m = 2, T = \{\{1, 2\}, \{1, 3\}\}$:

$$\begin{array}{llllllll} \text{Min} & p_1w_1 + p_2w_2 + p_3w_3 + p_{12}w_{12} + p_{13}w_{13} \\ \text{st} \\ & w_1 & & & & & \ge 1 \\ & & w_2 & & & & \ge 1 \\ & & & w_3 & & & \ge 1 \\ & w_1 & +w_2 & & +w_{12} & & \ge 1 \\ & w_1 & & +w_3 & & +w_{13} & \ge 1. \end{array}$$

In what follows a subset $S \subseteq \mathcal{V}$ is called w-tight if it satisfies the feasibility constraint for problem (8) with equality for a given basic feasible solution w, that is $w_T(S) = 1$.

Lemma 4. *Let $S \subseteq \mathcal{V}$ denote a w-tight subset. Then the subgraph induced on S by the spanning tree T is connected.*

Proof. Let $S_1, \ldots, S_k$ denote the connected components of the subgraph induced on S by the spanning tree T.

Feasibility conditions ensure that $w_T(S_i) \ge 1$ for all i, $1 \le i \le k$. The identity $w_T(S) = \sum_{i=1}^k w_T(S_i)$ then implies $w_T(S) \ge k$, which forces $k = 1$ because the subset S has unitary weight by assumption. □

The following lemma shows that the family of *w-tight* subsets is closed with respect to the union and the intersection.

Lemma 5. *Let $S_1, S_2 \subseteq V$ denote two w-tight subsets such that $S_1 \cap S_2 \neq \emptyset$. Then their intersection $S_1 \cap S_2$ as well as their union $S_1 \cup S_2$ are w-tight.*

Proof. Let us remark that there are no edges in T between the two subsets $S_1 \setminus S_2$ and $S_2 \setminus S_1$, because the subgraphs induced on S_1 and S_2 by the spanning tree T are connected by Lemma 6 and the spanning tree T contains no circuits. Thus the identity $w_T(S_1 \cap S_2) + w_T(S_1 \cup S_2) = w_T(S_1) + w_T(S_2)$ holds.

Since feasibility conditions ensure that $w_T(S_1 \cap S_2) \geq 1$ and $w_T(S_1 \cup S_2) \geq 1$ and because $w_T(S_1) = w_T(S_2) = 1$ by assumption, the above identity yields $w_T(S_1 \cap S_2) = 1$ and $w_T(S_1 \cup S_2) = 1$. □

Lemma 6. *Let $\Psi = \{S_u \subseteq V | S_u \ni u, w_T(S_u) = 1\}$ denote the collection of w-tight subsets containing vertex $u \in V$ and let $\Theta_u = \bigcap_{S_u \in \Psi} S_u$. Then the subset Θ_u is w-tight, that is Θ_u is the minimal element of the collection Ψ.*

Proof. Let us first remark that the collection Ψ is nonempty because for a given basic feasible solution w of problem (8) there exists at least one basic set containing u by Remark 2. The statement of the lemma then follows from Lemma 5. □

Lemma 7. *Let $w = (w_\gamma)_{\gamma \in \Gamma_T}$ denote an optimal solution of problem (8). Then $w_u = 1$ for all vertices $u \in \mathcal{V}$.*

Proof. Without loss of generality we can assume that the vector w maximizes the quantity $\sum_{\gamma \in \mathcal{V}} w_\gamma$ among all optimal solutions of problem (8).

Proving the lemma is equivalent to showing that the subset Θ_u as that of Lemma 6 reduces to the singleton $\{u\}$ for all vertices $u \in \mathcal{V}$. We will prove the statement by contradiction.

Let us assume that $\Theta_u \supsetneq \{u\}$, that is let $w_u > 1$.

The subset Θ_u is w-tight by Lemma 6, and thus connected by Lemma 4, therefore there exists a vertex $v \in \Theta_u$ such that the removal of edge $\{u, v\} \in T$ disconnects subset Θ_u in, say, subsets Θ_u^1 and Θ_u^2, where $u \in \Theta_u^1$, $v \in \Theta_u^2$. In particular let us observe that the inequalities $w_T(\Theta_u^1) \geq 1$ and $w_T(\Theta_u^2) \geq 1$ must hold by feasibility but that the subset Θ_u^1 cannot be w-tight, because the assumption $w_T(\Theta_u^1) = 1$ would contradict the minimality of subset Θ_u in the family of subsets Ψ. Thus the identity

$$w_{u,v} = w_T(\Theta_u) - w_T(\Theta_u^1) - w_T(\Theta_u^2) = 1 - w_T(\Theta_u^1) - w_T(\Theta_u^2)$$

implies $w_{u,v} < -1$.

Let $\varepsilon = \min\{\{w_T(S) - 1 | u \in S, w_T(S) > 1\}, -1 - w_{u,v}\} > 0$ and let w^ε denote the vector with components

$$\begin{cases} w_i^\varepsilon = w_i & \text{for } i \in \mathcal{V} \setminus \{u\} \\ w_u^\varepsilon = w_u - \varepsilon & \\ w_{i,j}^\varepsilon = w_{i,j} & \text{for } \{i, j\} \in T, \{i, j\} \neq \{u, v\} \\ w_{u,v}^\varepsilon = w_{u,v} + \varepsilon. & \end{cases}$$

The vector w^ε is still feasible for problem (8), because for all subsets $S \subseteq \mathcal{V}$ the inequality $w_T^\varepsilon(S) \geq 1$ holds. Consider the following cases.

Case 1: $u \notin S$. If node u does not belong to S the weight assignments with respect to vectors w and w^ε coincide, therefore $w_T^\varepsilon(S) = w_T(S)$, and $w_T^\varepsilon(S) \geq 1$ since $w_T(S) \geq 1$.

Case 2: $u \in S, v \notin S$. If node u belongs to S but the edge $\{u, v\}$ is not in $\mathcal{E}(S)$ the weight assignments with respect to vectors w and w^ε differ only on vertex u, therefore $w_T^\varepsilon(S) = w_T(S) - \varepsilon$. The subset S cannot be w-tight because otherwise the subset S would be a member of the collection Ψ, which in turn would yield $S \supseteq \Theta_u$ and $S \ni v$. Thus the inequality $w_T^\varepsilon(S) \geq 1$ still holds by the definition of the parameter ε.

Case 3: $u, v \in S$. By construction the weight assignments with respect to vectors w and w^ε differ only on node u and edge $\{u, v\}$, but $w_v^\varepsilon + w_{u,v}^\varepsilon = w_v - \varepsilon + w_{u,v} + \varepsilon = w_v + w_{u,v}$, therefore $w_T^\varepsilon(S) = w_T(S)$, and $w_T^\varepsilon(S) \geq 1$ since $w_T(S) \geq 1$.

Let us now evaluate the difference $z(w) - z(w^\varepsilon)$ given by the expression

$$z(w) - z(w^\varepsilon) = \sum_{i \in \mathcal{V}} (w_i - w_i^\varepsilon) p_i + \sum_{\{i,j\} \in T} (w_{i,j} - w_{i,j}^\varepsilon) p_{i,j} = \varepsilon(p_u - p_{u,v}).$$

Since the inequality $p_u \geq p_{u,v}$ trivially holds, the quantity $z(w) - z(w^\varepsilon)$ is nonnegative, which in turn implies that the vector w^ε is itself an optimal solution of problem (8). A contradiction then arises since the inequality

$\sum_{\gamma\in\mathcal{V}} w_\gamma < \sum_{\gamma\in\mathcal{V}} w^\varepsilon_\gamma$ holds and the vector w is selected as the optimal solution of problem (8) maximizing the sum of the node weights.

We can thus conclude that $w_u = 1$ for all vertices $u \in \mathcal{V}$. □

Lemma 8. *Let $w = (w_\gamma)_{\gamma\in\Gamma_T}$ denote an optimal solution of problem (8). Then $w_{u,v} = -1$ for all edges $\{u, v\} \in T$.*

Proof. Let us first remark that by Lemma 7 we can assume $w_u = 1$ for all vertices $u \in \mathcal{V}$.

Due to the feasibility requirement the inequality $w_T(\{i, j\}) \geq 1$ holds for all edges $\{i, j\} \in T$, but $w_T(\{i, j\}) = w_i + w_j + w_{i,j} = 2 + w_{i,j}$, therefore

$$w_{i,j} \geq -1 \tag{I}$$

for all edges $\{i, j\} \in T$.

Let $S_{i,j}$ be a basic set containing edge $\{i, j\} \in T$. The w-tight subset $S_{i,j}$ is connected by Lemma 4, and the edge $\{i, j\}$ disconnects subset $S_{i,j}$ in, say, subsets $S^1_{i,j}$ and $S^2_{i,j}$, where $i \in S^1_{i,j}$, $j \in S^2_{i,j}$. In particular let us observe that the inequalities $w_T(S^1_{i,j}) \geq 1$ and $w_T(S^2_{i,j}) \geq 1$ must hold by feasibility. Thus the identity $w_{i,j} = w(S_{i,j}) - w(S^1_{i,j}) - w(S^2_{i,j}) = 1 - w(S^1_{i,j}) - w(S^2_{i,j})$ implies

$$w_{i,j} \leq -1 \tag{II}$$

for all edges $\{i, j\} \in T$.

Inequalities (I) and (II) then yield $w_{i,j} = -1$ for all edges $\{i, j\} \in T$. □

Theorem 9. *The optimal objective function value of problem (8) is given by*

$$\sum_{i\in\mathcal{V}} p_i - \sum_{\{i,j\}\in T} p_{i,j}.$$

Proof. The assertion follows from Lemmas 7 and 8. □

Acknowledgements

The author wishes to deeply thank Prof. E. Boros for his continued support, and her daughter Sofia for inspiration.

References

[1] D. Hunter, An upper bound for the probability of the union, J. Appl. Probab. 30 (1975) 597–603.
[2] L. Lovász, M.D. Plummer, Matching theory, Ann. Discrete Math. 29 (1986) 307.
[3] A. Prékopa, B. Vizvári, G. Regös, Lower and upper bounds on probabilities of boolean functions of events, Rutcor Research Report 36-95, 1995.